



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/844,040	04/27/2001	Brian Seong-Gon Kim	089984-0276267	2687

27498 7590 12/22/2004

PILLSBURY WINTHROP LLP
2475 HANOVER STREET
PALO ALTO, CA 94304-1114

EXAMINER

CHEN, CHONGSHAN

ART UNIT	PAPER NUMBER
----------	--------------

2162

DATE MAILED: 12/22/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/844,040

Applicant(s)

KIM ET AL.

Examiner

Chongshan Chen

Art Unit

2162

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 22 June 2004.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-16 and 18-39 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-16 and 18-39 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

1. This action is responsive to communications filed on 22 June 2004. Claims 1-16 and 18-39 are pending in this Office Action. Claim 17 is canceled.

Response to Arguments

2. Applicant's arguments filed on 22 June 2004 have been fully considered but they are not persuasive.

3. As per applicant's arguments regarding "the 'enough documents' type phrases of claims 2, 20 and 26, Applicants' specification as filed at page 12, last three paragraphs, discloses the details of how one embodiment makes the decision of whether there are enough or not enough documents ... 112, second paragraph, rejection of claims 2, 20 and 26 should be withdrawn" have been considered but are not persuasive. The supporting paragraphs disclose using total weight value to determine whether there are enough or not enough documents. However, the claims do not show using the total weight for the determining process. Therefore, the arguments are not persuasive and the 112 second paragraph rejection is maintained.

4. In response to applicant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, Ferguson teaches a document management system extracts key words from documents and categorizes documents based on the extracted key words (Ferguson, col. 8, lines

Art Unit: 2162

12-21). Ho teaches extracting phrases from documents. The phrases extractor ignores negative/non-important words, and replaces non-root words with their roots (Ho, col. 2, lines 33-36, col. 6, lines 50-55, col. 12, lines 22-26). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the document categorizing system of Ferguson by replacing its keyword extractor with the phrase extractor of Ho. It is because the negative words are meaningless for the categorizing purpose, and the non-root keywords make it difficult for the document categorizer to compare the words. The ordinary skilled artisan would have been motivated to ignore the negative words and replace the non-root words with their roots in order to make the document categorizer more accurate and efficient.

5. In response to applicant's argument that the examiner's conclusion of obviousness is based upon improper hindsight reasoning, it must be recognized that any judgment on obviousness is in a sense necessarily a reconstruction based upon hindsight reasoning. But so long as it takes into account only knowledge which was within the level of ordinary skill at the time the claimed invention was made, and does not include knowledge gleaned only from the applicant's disclosure, such a reconstruction is proper. See *In re McLaughlin*, 443 F.2d 1392, 170 USPQ 209 (CCPA 1971).

6. As per applicant's arguments regarding the references do not teaches assigning documents in the preprocessed collection of documents to category have been considered but are not persuasive. Ho teaches preprocessing the document by ignore the negative words and changing the non-root words to their roots (Ho, col. 2, lines 33-36, col. 6, lines 50-55, col. 12, lines 22-26). Furthermore, Ho teaches assigning document in the preprocessed collection of document to a category (Ho, col. 2, lines 33-36, "phrases in the document are automatically

Art Unit: 2162

extracted based on grammar and dictionaries. From these phrases, categories in category hierarchy are identified, and the document is linked to the categories”).

7. As per applicant's arguments regarding the combination of two references would render the Ferguson search function incapable of performing in its intended fashion have been considered but are not persuasive. The document management system of Ferguson includes several functions including categorizing, searching, etc. The claimed invention only claims document categorizing, not searching. The Ferguson and Ho's combined system for document categorizing performs the claimed categorizing method. Therefore, the arguments are not persuasive.

8. As per applicant's arguments regarding “applicant’s preprocessing methodology is patentably distinct from Ho because Applicant’s invention does not use a hashing function to convert words to root form. Rather, application’s invention uses one or more dictionary-type look-up or comparison files stored within the system to preprocess the document” have been considered but are not persuasive. It is noted that the features upon which applicant relies (i.e., uses one or more dictionary-type look-up or comparison files stored within the system to preprocess the document) are not recited in the rejected claim(s).

9. As per applicant's arguments regarding the claim 2 the references do not teach the testing step, the constructing step or the assigning step of Applicants’ claimed invention have been considered but are not persuasive. Ho teaches testing to determining if there are enough documents (Ho, col. 9, line 64 – col. 10, line 1, when the number of documents linked to it exceeds a predetermined value), constructing the new category (Ho, col. 9, line 64 – col. 10, line

Art Unit: 2162

1, a new category will be created) and assigning the seed document to a category (Ho, col. 9, line 64 – col. 10, line 1). Therefore, the arguments are not persuasive.

10. As per applicant's arguments regarding claims 3-7 the references do not teach removing the punctuation marks, replaces upper-case characters with lower-case characters, replaces the non-root words with root words and removes the articles from the string of characters have been considered but are not persuasive. Ho teaches extracting phrases from documents for categorizing purposes. The phrases extractor ignores negative/non-important words, and replaces non-root words with their roots (Ho, col. 2, lines 33-36, col. 6, lines 50-55, col. 12, lines 22-26). It is because the negative words are meaningless for the categorizing purpose, and the non-root keywords make it difficult for the document categorizer to compare the words.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to ignore the negative words including removing the punctuation marks and removes the non-important articles from the string of characters, replaces upper-case characters with lower-case characters and replace the non-root words with their roots in order to make the document categorizer more accurate and efficient.

11. As per applicant's arguments regarding claim 20 the references do not teach the use of a temporary category have been considered but are not persuasive. The applicants state the temporary category as claimed does not have category properties assigned. It is merely a holding area, or testing area (argument file on 6/22/2004, page 14, second complete paragraph).

Ferguson teaches the temporary category (folder) that is not associated with any category (Ferguson, col. 7, lines 9-13). The folder is merely a holding area for documents.

Art Unit: 2162

12. As per applicant's arguments regarding claims 23-25, the references do not teach the use of anchor-text character string have been considered but are not persuasive. The applicants' admitted prior art discloses the use of anchor-text (specification, page 14, last paragraph – page 15, line 2). Therefore, the arguments are not persuasive.

Claim Rejections - 35 USC § 112

13. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

14. Claims 2, 20 and 26 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

The phrase “determine if there are enough documents”, “enough documents” and so on are indefinite. It is unclear how many documents are enough documents?

Claim Rejections - 35 USC § 103

15. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

16. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any

Art Unit: 2162

evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

17. Claims 1-16 and 18-39 are rejected under 35 U.S.C. 103(a) as being unpatentable over Ferguson et al. ("Ferguson", 6,237,011) in view of Ho et al. ("Ho", 6,571,240).

As per claim 1, Ferguson teaches a method of categorizing an initial collection of documents, each document being represented by a string of characters, the method comprising the steps of:

identifying predefined characters in the string of characters from the documents in the initial collection of documents to form identified characters (Ferguson, col. 8, lines 12-21, "... the key words and/or attributes are automatically extracted ...");

constructing a number of categories from the string of characters of the preprocessed collection of documents (Ferguson, col. 8, lines 12-32); and

assigning each document in the preprocessed collection of documents to a category to form a hierarchy of categories of documents (Ferguson, col. 8, lines 12-32, "categorizing a document into one or more categories ...").

Ferguson does not explicitly disclose changing the identified characters in the documents in the initial collection of documents to form a preprocessed collection of documents, each of the preprocessed collection of documents represented by a preprocessed string of characters. Ho teaches changing the identified characters in the documents (Ho, col. 12, lines 16-26, ignores negative words and changes non-roots words to their roots). Therefore, it would have been

Art Unit: 2162

obvious to one of ordinary skill in the art at the time the invention was made to modify the document categorizing system of Ferguson by incorporating the limitation of changing the identified characters in the documents in the conventional manner as disclosed by Ho (Ho, col. 2, lines 33-36, col. 6, lines 50-55, col. 12, lines 22-26). It is because the negative words are meaningless for the categorizing purpose, and the non-root keywords make it difficult for the document categorizer to compare the words. The ordinary skilled artisan would have been motivated to ignore the negative words and replace the non-root words with their roots in order to make the document categorizer more accurate and efficient.

As per claim 2, Ferguson and Ho teach all the claimed subject matters as discussed in claim 1, and further teach

clearing a temporary category and selecting a seed document from the preprocessed collection of documents as a first document of the temporary category (Ferguson, col. 8, lines 12-32);

collecting documents from the preprocessed collection of documents that are similar to the seed document into the temporary category (Ferguson, col. 8, lines 12-32);

testing to determine if there are enough documents in the temporary category to merit construction of a new category (Ho, col. 9, lines 64-67);

constructing the new category from the temporary category and generating a heading for the new category if there are enough documents in the temporary category to merit construction and generation (Ho, col. 9, lines 64-67);

Art Unit: 2162

assigning the seed document to a category reserved for documents not belonging to any specific category if there are not enough documents in the temporary category (Ferguson, col. 8, lines 12-32, Ho, col. 9, lines 64-67); and

marking the documents assigned to any category in the preprocessed collection of documents as processed (Ferguson, col. 8, lines 12-32).

As per claim 3, Ferguson and Ho teach all the claimed subject matters as discussed in claim 2, except for explicitly disclosing the predefined characters include punctuation marks, and the changing step removes the punctuation marks from the string of characters. However, Ho teaches extracting phrases from documents for categorizing purposes. The phrases extractor ignores negative/non-important words, and replaces non-root words with their roots (Ho, col. 2, lines 33-36, col. 6, lines 50-55, col. 12, lines 22-26). It is because the negative words are meaningless for the categorizing purpose, and the non-root keywords make it difficult for the document categorizer to compare the words. Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the Ferguson and Ho's combined categorizing system by ignoring the negative words including removing the punctuation marks and removes the non-important articles from the string of characters, replaces upper-case characters with lower-case characters and replace the non-root words with their roots. The motivation being to make the document categorizer more accurate and efficient.

Claims 4-7 are rejected on grounds corresponding to the reasons given above for claim 3.

As per claim 8, Ferguson and Ho teach all the claimed subject matters as discussed in claim 2, and further teach loading a character string from the seed document into a memory

Art Unit: 2162

location to initialize the values of a number of category properties for the temporary category (Ferguson, col. 8, lines 12-32).

As per claim 9, Ferguson and Ho teach all the claimed subject matters as discussed in claim 8, and further teach

determining if there are documents in the preprocessed collection of documents that have not been processed with respect to the temporary category (Ho, col. 9, lines 64-67);

if there are documents in the preprocessed collection of documents that have not been processed with respect to the temporary category, selecting a next document from the preprocessed collection of documents and measuring a similarity of the preprocessed string of characters of the text document using a similarity test between the next document and the values of the number of current category properties (Ferguson, col. 8, lines 12-32);

including the next document in the temporary category if the next document passes the similarity test (Ferguson, col. 8, lines 12-32);

updating the values of the number of category properties of the temporary category when the next document is included (Ferguson, col. 8, lines 12-32); and

rejecting the next document if the next document fails the similarity test (Ferguson, col. 8, lines 12-32).

As per claim 10, Ferguson and Ho teach all the claimed subject matters as discussed in claim 9, and further teach repeating the steps of claim 9 for all documents in preprocessed collection of documents (Ferguson, col. 8, lines 12-32).

As per claim 11, Ferguson and Ho teach all the claimed subject matters as discussed in claim 2, and further teach collecting more similar documents from a number of existing categories (Ferguson, col. 8, lines 12-32).

As per claim 12, Ferguson and Ho teach all the claimed subject matters as discussed in claim 11, and further teach

determining if there are more documents in a number of existing categories that have not been processed with respect to the temporary category (Ferguson, col. 8, lines 12-32);

if there are documents in the number of existing categories that have not been processed with respect to the temporary category, selecting a next document from the number of existing categories as a selected document and measuring a similarity of the preprocessed string of characters of the selected document using a similarity test between the selected document and values of a number of current category properties (Ferguson, col. 8, lines 12-32);

including the selected document in the temporary category if the selected document passes the similarity test (Ferguson, col. 8, lines 12-32); and

rejecting the selected document if the selected document fails the similarity test (Ferguson, col. 8, lines 12-32).

As per claim 13, Ferguson and Ho teach all the claimed subject matters as discussed in claim 12, and further teach repeating the steps of claim 12 for all documents in the number of existing categories (Ferguson, col. 8, lines 12-32).

As per claim 14, Ferguson and Ho teach all the claimed subject matters as discussed in claim 8, and further teach the category properties includes a string of characters selected from the group consisting of a longest common substring in the title, a longest common substring in the

Art Unit: 2162

body; and a document type index measured as list of fractional numbers for each document type (Ferguson, col. 8, lines 12-32).

As per claim 15, Ferguson and Ho teach all the claimed subject matters as discussed in claim 14, and further teach categorizing documents into categories (Ferguson, col. 8, lines 12-32), the documents inherently includes news article, technical documents, and poems.

As per claim 16, Ferguson and Ho teach all the claimed subject matters as discussed in claim 2, and further teach making sub-categories if there are too many documents in a given category; and post-processing the number of categorized lists of documents (Ho, col. 9, line 62 - col. 10, line 9).

As per claim 18, Ferguson and Ho teach all the claimed subject matters as discussed in claim 2, and further teach the seed document is a first document in the preprocessed collection of documents (Ferguson, col. 8, lines 12-32).

As per claim 19, Ferguson and Ho teach all the claimed subject matters as discussed in claim 2, and further teach the seed document is a document with a highest rank value among the documents not marked as processed in the preprocessed collection of documents (Ferguson, col. 8, lines 12-32).

As per claim 20, Ferguson and Ho teach all the claimed subject matters as discussed in claim 2, and further teach the temporary category is tested to determine if there are enough documents in the temporary category to merit construction of a new category by accumulating the weight of each document when each document can contribute uniform weight or different weight based on the rank value of each document with higher ranked document given more weight (Ferguson, col. 8, lines 12-32).

As per claim 21, Ferguson and Ho teach all the claimed subject matters as discussed in claim 2, and further teach except for explicitly disclosing the heading is a longest common substring in a title. It would have been obvious to one of ordinary skill in the art at the time the invention was made to use the longest common substring in a title as category heading because the longest common phrase in the title best describes the topic of the category. Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the Ferguson and Ho's combined system by using the longest common substring in a title as category heading because the longest common phrase in the title best describes the topic of the category. The motivation being to use the longest common phrase to better describe the category.

Claim 22 is rejected on grounds corresponding to the reasons given above for claim 21.

As per claim 23, Ferguson and Ho teach all the claimed subject matters as discussed in claim 1, except for explicitly disclosing determining if an anchor-text character string is available for the documents in the initial collection of documents; and attaching an anchor-text character string to the string of characters that represents the documents in the initial collection of documents when the anchor text character string is available. However, it is well known in the art determining if an anchor-text character string is available for the documents in the initial collection of documents; and attaching an anchor-text character string to the string of characters that represents the documents in the initial collection of documents when the anchor text character string is available (applicants admitted prior art, specification, page 14, last paragraph – col. 15, line 2). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the Ferguson and Ho's combined system by

Art Unit: 2162

incorporating the use of anchor text as admitted by the applicants. The motivation being to help determine the relevancy of the page to a given category.

As per claim 24, Ferguson and Ho teach all the claimed subject matters as discussed in claim 23, except for explicitly disclosing the anchor-text character string is a text used most frequently by hypertext documents. However, it is well known in the art that the anchor-text character string is a text used most frequently by hypertext documents (applicants admitted prior art, specification, page 14, last paragraph – col. 15, line 2). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the Ferguson and Ho's combined system by incorporating the use of anchor text as admitted by the applicants. The motivation being to help determine the relevancy of the page to a given category.

As per claim 25, Ferguson and Ho teach all the claimed subject matters as discussed in claim 23, except for explicitly disclosing the anchor-text character string is a text with a highest partial extrinsic rank value. However, it is well known in the art that the anchor-text character string is a text with a highest partial extrinsic rank value (applicants admitted prior art, specification, page 14, last paragraph – col. 15, line 2). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the Ferguson and Ho's combined system by incorporating the use of anchor text as admitted by the applicants. The motivation being to help determine the relevancy of the page to a given category.

Claims 26-39 are rejected on grounds corresponding to the reasons given above for claims 1-16 and 18-25.

Conclusion

18. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Towell (6,052,680) discloses a preprocessor that removes the punctuation marks, replaces upper-case characters with lower-case characters and removes the articles from the string of characters (Towell, col. 5, lines 30-43).

Mehrle (5,794,236) discloses a computer-based system for classifying documents into a hierarchy and linking the classifications to the hierarchy.

19. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Art Unit: 2162


Contact Information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Chongshan Chen whose telephone number is (571)272-4031. The examiner can normally be reached on Monday - Friday (8:00 am - 4:30 pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John E Breene can be reached on (571)272-4107. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Chongshan Chen
December 11, 2004


JEAN M. CORRIELUS
PRIMARY EXAMINER